National Research
FLAGSHIPS
Wealth from Oceans

CSIRO

Chesapeake Bay Blue Crab Stock Assessment Review 2011

Baltimore, Maryland
29-31 March 2011

www.csiro.au

Dr Cathy Dichmont
CSIRO Wealth from Oceans Flagship
Ecoscience Precinct, 41 Boggo Road, Dutton Park Qld 4102
AUSTRALIA

Representing the Center for Independent Experts

Contents

Executive Summary

The review workshop for the Chesapeake Bay blue crab assessment took place in Baltimore Maryland on March 29-31, 2011. In attendance were review panel members Drs Dichmont, Addison and Ernst, crab stock assessment authors, other scientists involved in the stock assessment, and fishery managers. The review was undertaken in a very co-operative light with reasonable requests for additional work met, including providing more diagnostic plots and a sensitivity test. The panel members were presented with information and reports on basic biological rates and how they were derived, the data inputs to the model, the main assessment, reference points and stock status relative to the reference points. Two reports on additional assessment models were also provided, thereby addressing the sensitivity of the crab assessment to model structure. This information correlated well with the review Terms of Reference, which is appreciated.

Given the three assessments, the main assessment – the sex-specific catch multiple survey analysis (SSCMSA) – is an improvement on past models and the best of the three models submitted for review. The SSCMSA is a two-stage sex disaggregated model that uses three indices of abundance from the surveys, one of which (the winter dredge survey) is treated as absolute. The two stages are year 0 and year 1+ animals. A fundamentally positive development of this model is the integration of the reference point calculations internal to the model using per recruit analyses. This means that assumptions in the assessment can be consistently carried into the reference point calculations. This is a major advance and is highly supported.

This model also differs from its predecessors in that it models male and female populations separately. In light of the recent major changes to the fishery management that have changed the sex ratio of the catch, this is an appropriate advance on previous models.

In order for this integration of incorporating reference points into the model to take place, the model has to incorporate a renewal function, in this case a modified form of the Ricker stock recruitment relationship. This form of the relationship is not plausible in that it assumes very high productivity if females are present, even for a zero male population size i.e. there is no sperm limitation in the population. There is no information on this lower end of the stock-recruitment curve to support such a relationship assumption. A sensitivity test using the traditional approach and a dispensation form was requested but there was not enough time for the assessment team to complete this request given the work required. This is an important aspect of the model that should be further justified given the precautionary approach.

The derivation of the reference point calculations is correctly implemented in the model. The basis of choosing reference points traditionally (at least in other countries' management systems) is not only the domain of the assessment scientists. In this context, the report should provide more choices of reference points, especially the target reference points. Even so, the use of both the Federal reference points system to define the limit reference points and the regional Management Councils' system to define a target reference point is an improvement. However, the assessment team has chosen a limit reference point value that is not in the precautionary part of the potential range as argued.

By international standards, the fishery is considered data rich, especially in terms of the amount and extent of surveys. Some of these data are not included in the assessment - comments regarding these excluded data and which data to assign high priority to are included in the recommendations. Despite the quantum of data, there is uncertainty with respect to the catch and effort data, with the effort data is not used in the model. Although the analyses of adjusted catch data to under-reporting seems reasonable, there is a need to independently verify the assumed values, given that the effect on the catch series is large. Even treating this as a sensitivity test (by using the raw data or a mid range between raw and adjusted) in the model would be an improvement.

A test run during the review highlighted that the assessment model output is extremely sensitive to whether the winter dredge survey is assumed to be a relative or absolute index of abundance. This is in terms of the absolute stock status, reference points and stock status relative to the reference points. Using the relative index model, the likelihood of the model fit is very much lower because the fit to the data is better (despite estimating more parameters). This SCMSA model in its current configuration is therefore not robust to this assumption. This is of great concern, as traditionally surveys are not treated as absolute abundance indices. It is highly recommended that: a) a sensitivity test is run that uses the raw winter dredge survey indices (i.e. not converted to absolute), and b) the derivation of the catchability of the winter dredge survey index of abundance should be checked in detail and possible sources of this difference between input and estimated winter dredge catchabilities (both assessment and survey) investigated.

This assessment is a valid approach and an improvement on the previous assessments and therefore should be adopted as the basis for management advice. However, the justification for treating the winter dredge survey as an absolute index of abundance needs to be substantiated given that the best model fit is achieved by treating the survey as a relative index of abundance (which is the more widely accepted method) and this has important implications to management. This is the highest short-term priority.

Background

Blue crab is the most important commercial fishery in Chesapeake Bay. The blue crab stock has been subject to Baywide stock assessments on two previous occasions. In the years between benchmark assessments, updates on the stock status are provided by the NOAA Chesapeake Bay Office's Chesapeake Bay Stock Assessment Committee.

The first Baywide stock assessment (Rugulo *et al*., 1997) was conducted using a length-based approach to estimate exploitation, and an unweighted average of the four principal fishery-independent surveys to determine abundance. Consequently biological reference points were crude.

In 2001, the technical subcommittee of the Bi-State Blue Crab Advisory Committee (BBCAC) developed a new management framework that relied on exploitation and biomass threshold and target reference points. Threshold reference points were proposed based on maintaining 10% of the virgin spawning potential and on the lowest observed abundance in the surveys. A target exploitation rate that would lead to an effective doubling of the spawning stock present in 2001 was also selected. The most recent Baywide benchmark assessment for blue crab in the Chesapeake Bay was conducted in 2005 (Miller *et al*., 2005). This assessment critically evaluated and revised estimates of the natural mortality rate, the impact of reporting changes on landings estimates, and spawning potential ratio reference points. The 2005 assessment recommended adopting the exploitation fraction, defined as the proportion of crabs available at the beginning of the season that are subsequently harvested, in place of less intuitive measures (F) used in previous assessments. Estimates of exploitation fractions were calculated based on the Baywide winter dredge survey (WDS) and within a modified catch-survey analysis that permitted the use of multiple surveys' data. The approach used in the 2005 assessment was reviewed by a panel of international scientists with expertise in crustacean fisheries who found that it was a substantial improvement over previous assessments (Haddon, 2005). However, the panel also identified issues to be addressed in future assessments. In particular, the panel recommended exploration of the impact of density-dependent processes in life history traits, improvements to the fishery-independent surveys, particularly with regard to catchability, the possibility of developing a sex-specific assessment model and reference points, and a fuller analysis of the impacts of uncertainty on all aspects of the assessment.

Since the 2005 assessment, the three management jurisdictions have implemented a range of regulatory changes aimed at attaining the target exploitation rate of 46% of the available stock. In 2009 the NOAA Chesapeake Bay Office initiated and supported development of another benchmark assessment and that assessment (Miller *et al*., 2011) is the basis of this CIE Peer Review.

Description of Review Activities

The review workshop for the Chesapeake Bay blue crab assessment took place in Baltimore Maryland on March 29-31, 2011. On March 28 there was a Conference by the blue crab advanced research consortium which provided interesting background to the fishery but was not essential to the reviewers. The NOAA Chesapeake Bay Office provided the documents on a website for the Review (Appendix 1). After a request for the code in a form that could be run in ADMB and R, these were provided to the review team via the website.

The objective of the CIE Chesapeake blue crab review was to review and discuss input data, assumptions and parameters to the model, the assessment itself and the biological reference points for the population.

The Terms of Reference were:

a) Critically assess, and where necessary revise, the life history and vital rates of blue crab in the Chesapeake Bay that are relevant to an assessment of the stock.
b) Evaluate and recommend biological reference points for the Chesapeake Bay blue crab population. The potential for implementing sex-specific reference points should be evaluated.
c) Describe and quantify patterns in fishery-independent surveys. Analyses should include an evaluation of the impacts of environmental and abiotic factors on survey catches, to maximize the information content of resultant survey time series.
d) Describe and quantify patterns in catch, effort and survey-based estimates of exploitation by sector and region, including analyses that examine the impacts of reporting changes and trends in CPUE.
e) Develop and implement assessment models for the Chesapeake blue crab fisheries. In particular, models that permit estimates of the trends and status of the crab population and fisheries on a sex-specific basis should be evaluated.
f) Examine density-dependent exploitation patterns derived from survey-based and model-based approaches.
g) Characterize scientific uncertainty with respect to assessment inputs and stock status.

h) Evaluate stock status with respect to reference points.

The meeting was chaired by Dr Derek Orner of NOAA on behalf of the CIE review team, but the summary review was the view of the CIE panel members.

Dr Tom Miller gave a presentation on the first day of the review covering all aspects of the above Terms of Reference (ToR). On the second day, the review panel and assessment team discussed the research against each ToR. Upon request, the team provided additional documentation and information. The assessment team was extremely helpful in this review. The agenda for the review is provided in Appendix 2.

Additional information supplied:
1. Correction of the lagged plots provided in the report
2. Extra residual and predicted versus observed plots
3. Extra run with winter dredge survey as a relative index of abundance, and
4. Extra run using a standard Ricker stock-recruitment function.

Summary of Findings

In order to facilitate a more logical flow for the review, the order of the terms of reference has been slightly changed. This means that comments on the biological reference points are provided after those on the assessment model since these flow from the latter.

ToR: Review and Critically assess and where necessary revise the life history and vital rates of blue crab in the Chesapeake Bay that are relevant to an assessment of the stock.

Stock structure

Based on genetics studies and larval distribution it has been assumed that, from a stock assessment point of view, Chesapeake Bay blue crab is effectively a single stock. This is based on several studies. This assumption seems to be valid given the evidence (which is more than many fisheries making this assumption have had elsewhere).

Growth rates

There are several sources of growth rate information - being lipofuscin analyses, moult process models and stock-recruitment estimates from stock enhancement studies on both larvae, juvenile and small crabs. Growth was shown to be season dependent, which affects the timing of recruitment within the year by 10%. However, growth rates obtained from lipofuscin are not directly or indirectly used in the model other than through inference or body of evidence.

There are three separate processes that have an implicit growth rate:
1. separating the survey size frequency data into age 0 and 1+
2. partial recruitment - a parameter to implicitly link within year indices of abundance to key biological rates, and
3. maximum age and therefore natural mortality.

The survey age classes were derived from a cut off table based on space and time of the survey, however there are very clear modes between age zero and 1+ animals from the graphs provided during the review.

Lipofuscin analysis is used as a method to age crab in Chesapeake Bay. This information is influential. Although there has been some criticism of this method in other parts of the world, it seems appropriately applied here although there could be a question over whether it is a chronometer in oxidative stress areas? In this regard, the ages are provided with a 6 month error range which is therefore appropriate. Furthermore, the sample size of aged crabs is very large.

> **Recommendation 1.** It is unclear whether these different inputs of implicit age are internally consistent. Some evidence of this should be provided.

Moult to maturity

Circumstantial evidence and limited physiological evidence that, for example, the Y organ does not degenerate, point to blue crabs having terminal moult. This assumption, which is influential in the data analysis and model structure, seems appropriate.

Age and size at maturity

Aquaculture studies have shown that females can mature within the first year under ideal conditions. However, the authors argue that given the temperatures during the year and typical megalopial settlement dates, it is unlikely that crabs mature within their first year in Chesapeake Bay. This seems to be an appropriate assumption.

Mating and spawning periods

Female blue crabs are only receptive immediately after moulting. There is intense competition among males during this receptive period. Mating mainly occurs from May to October. In upper Chesapeake Bay none of the upper bay inseminated females will produce broods in the same year as mating whereas an unknown fraction in the lower bay could do so. In the lower Chesapeake Bay, there is evidence that some unknown fraction of females inseminated can release larvae in the same period.

Fecundity

Prager et al. (1990) showed a linear relationship between fecundity and carapace width with r squared of 0.24. Wells (2009) re-examined this relationship for the low-density period and found that there was a significant decrease in size of mature crab from the 1980's to 2005. In recent studies, there is little or no evidence for a linear relationship. Recent, but still preliminary data, show a change in fecundity as a function carapace width occurring from empirical evidence.

> **Recommendation 2.** A possible change in fecundity with respect to size should be monitored in terms of its impact on the assessment.

Dickinson et al. (2006) showed that average sized crabs spawned at the beginning of the season and produce eight clutches within a full 25 week spawning season. Larger animals produce larger clutches, less frequently than smaller animals. Darnell et al. (2009) concluded that the majority of reproductive output of individual females comes from a few initial broods in a female's lifetime.

Recommendation 3. In the assessment model, fecundity is not weighted by age, whereas it is likely to be a consideration and quite a major factor within the assessment where the category age 1+ implies that all animals greater than age 1 are equally fecund even though 2+ animals are unlikely to breed. Splitting the model into three age classes should be considered.

Recommendation 4. A research task for future assessments is to study age related fecundity and other related fecundity patterns for Chesapeake Bay.

Juveniles

Lipcius et al. (2007) showed that the juvenile period is critical for crabs with the state/condition of nursery grounds also being very important. There have been several studies on predation and nursery grounds. Etherington et al. (2003) showed that mortality rates on seagrass were equivalent to emigration rates. There seems to be key lower and upper Chesapeake Bay dynamics and these were not considered within the assessment. Given this and other strong spatial dynamics, it is likely that the assessment should ultimately have a stronger spatial element than it does presently.

Natural mortality

Natural mortality (M) is indirectly calculated using key biological rates, such as maximum age and growth parameters, and environmental factors rather than estimating M directly in the model. This seems appropriate and is the norm internationally.

Originally Rugolo et al. (1997) based his M estimate on a maximum age of eight. Miller at al. (2005) subsequently reviewed estimates of natural mortality for the previous assessment. The values of M used in the present assessment were based on work undertaken by Hewitt et al. (2007), which uses several sources of data to calculate M.

Brownie tag-return models in Lipcius and Smith (Assessment working paper 1) were used to calculate survivorship. However, the extensive tagging data from 2001-2009 (with respect to the assessment) are in essence, only used to calculate maximum age, which is then used to calculate a range of M values. This study therefore calculates M for female crabs that range from 0.6-0.8. The Base Case assessment used 0.9, with sensitivity tests at 0.6 and 1.2. Reference points were tested with the high and low M values as well.

A distribution of M (Fig 2.2 in the Assessment report) gives a range of M from 0.3 - 2.35. The use of 0.9 in the Base Case, rather than using the tagging data values for the Base Case, is due to the fact that the tagging data is of post moult females where mortality is likely to be higher than for the whole population. This assumption appears to be sound. Sensitivity tests to M were undertaken, which is appropriate.

However, the tagging data are underutilised for such an extensive data set. In reality, these data should be incorporated into the model. In addition, the form of the Brownie calculation is fairly basic and could be further improved so that it includes M and F, or M and catchability and effort.

Recommendation 5. In future, use the tagging data in the assessment to estimate female fishing mortality and natural mortality.

Recommendation 6. In the interim, modify the Brownie model to estimate survivorship and natural mortality through a parameterisation that includes M and F, or M and catchability and effort.

ToR: Describe and quantify patterns in fishery-independent surveys. Analyses should include an evaluation of the impacts of environmental and abiotic factors on survey catches, to maximize the information content of resultant survey time series.

Size at age conventions

Previous assessments have used size composition data from fishery independent surveys (FIS) to develop estimates of abundance. Bonzek (pers. comm.) explored the consequences of alternative demarcations of size-at-age vectors.

In this assessment, the size definitions of age classes have not been rigorously or fully evaluated, partly due to inconsistent reporting of size in some surveys. However, figures presented during the review showed reasonably clear modes in the data.

FIS time series

Three FIS are used in the assessment: 49 years of VIMS trawl survey sampling data (taken from the southern portion of the Bay), 28 years of MD trawl survey data (taken in the eastern shores of the Bay) and 16 years of winter dredge survey data (Baywide and treated as an absolute index of abundance).

Data from several other surveys were not used in the assessment because they were either too short or too localised. Potentially the most valuable of these are the Chesapeake Bay Multispecies Monitoring and Assessment Program survey data, even though they have only been gathered since 2003, as it is a Baywide survey. The argument that it is too short for use in the assessment is not supported given that there is a large amount of data already used in the assessment. This means that incorporation of the additional parameters would be supported.

Recommendation 7. Include the Baywide Multispecies Monitoring and Assessment Program in the assessment as a survey index of abundance.

Statistical analysis of FIS

The FIS time series are in the form of count data that are characterised by being zero inflated. Apart from the Virginia trawl survey (which was not standardised), a similar method was used to standardise the FIS indices of abundance. The use of the delta method is appropriate in this instance in that this method models the probability of presence and then models abundance given occurrence. A generalised linear model (GLM) is used to develop standardised indices of abundance. The best fitting model was developed using AIC. The model considers: Year, Month, Strata, Salinity, Temperature and Depth. For each model fit, variances of the indices were generated using jacknife estimates.

However, given the evidence of strong temporal and spatial structure in the data, it is clear that interaction terms in the GLM should be considered. Although interaction terms cannot be implemented in the package used, this should be modified or remodelled.

> **Recommendation 8.** Incorporate interaction terms in the second stage of the delta GLM, especially with respect to strata and year, temperature and year for the winter dredge survey.

In the delta GLM analysis, there is no link in the assessment report between the final models stated as used in the analyses and the findings shown in Tables 3.4 and 3.5. These were checked during the review and it was shown that they were correctly implemented and the text appears to be incorrect (see section on "Changes to the documentation" for further details).

Virginia juvenile finfish and blue crab trawl survey (since 1955)

There are seven spatial strata within the Virginia trawl surveys, which are undertaken monthly with both a fixed and random survey design. The analysis focused on three principle tributaries, because the spatial extent of the survey changed over time. Therefore only three of the strata are used as an index of abundance in the assessment as it was argued that the four extra strata were added later in the survey design. Furthermore, this index is not standardised. The assessment team rather calculated a geometric mean for each tributary and then took the average of these.

The argument of not including the extra strata is not supported given that there are years where a mechanism of comparison was included in the Virginia trawl survey expansion coverage. Either a standardisation process could have combined this data set or these extra strata could have been added as an additional survey.

> **Recommendation 9.** Include in the assessment model either using the four other Virginia trawl survey strata as an additional index of abundance or (better option) analyse the Virginia trawl survey and standardise into a single index.

MD DNR Trawl survey (since 1977)

This survey has inconsistent coverage both spatially and temporally. The survey is conducted from May to November and uses fixed sites. From 1989, size measurements were taken. The best model for age 0 abundance indices were developed using both design factors and temperature.

In Fig 3.9 of the assessment document, the two peaks in the modelled index have no corresponding observed data points. During the review, it was shown that the observations are outside the upper limit of the axes in the plots. This should be corrected in the document.

Winter dredge survey (used from 1991/2)

This is a key dataset as it is used in the assessment model as an absolute index of abundance. This is unusual given that most modellers are unable to accurately calculate survey catchability relative to the whole population (therefore treating survey indices as being a relative index of abundance). The survey is a stratified random design and then a subsequent adaptive design is used where additional sampling occurs in high-density areas.

The survey data have extensive spatial and temporal components, which appear to have not been used. See "Statistical analysis of FIS" section above.

The dredge efficiency figures are obtained from depletion analyses, which are undertaken in high-density areas for practical reasons. Absolute abundance is extrapolated to the whole area using a GIS estimated area. It should be noted that each jurisdiction uses different survey gear and different survey designs. The survey gear catch efficiency is estimated for each year, but not with respect to spatial distribution changes.

> **Recommendation 10.** Conduct more detailed analyses of the dredge survey catchability estimates, especially with regard to availability or else undertake a spatial model of the survey.
>
> **Recommendation 11.** Consider applying consistent methods across jurisdictions.
>
> **Recommendation 12.** The two-stage component of the dredge survey is an extremely valuable component of the analysis and should be continued.
>
> **Recommendation 13.** The dredge survey is essential and should be continued.

During the review, a sensitivity test was undertaken in which the assessment was modified so as to treat the winter dredge survey index as a relative index of abundance. This causes substantial changes to the model outputs – the model estimates much lower dredge survey catchability, fits the winter dredge survey index better and estimates of fishing mortality are higher than in the Base Case. Furthermore, this test has much better likelihood values than the Base Case, which treats the winter dredge survey as an absolute index of abundance. These results show that the assessment is extremely sensitive to this assumption, in terms of absolute stock status, reference points and stock status relative to reference points (see Appendix 3). This model is therefore not robust to this assumption, which is of great concern given that indices of abundance are traditionally not treated as an absolute index of abundance. In the review, the panel was not given detailed information about the winter dredge survey in order to clearly analyse the validity of this assumption.

> **Recommendation 14.** Run a sensitivity test that uses the raw standardised winter dredge survey indices (not converted to absolute).
>
> **Recommendation 15.** Review the derivation of the absolute index of abundance in detail and investigate possible sources of the difference between the absolute and relative model runs (both in terms of the assessment and the survey). This is the highest short-term priority.

ToR: Describe and quantify patterns in catch, effort and survey-based estimates of exploitation by sector and region, including analyses that examine the impacts of reporting changes and trends in CPUE.

Several gear types are used to catch blue crab – there is a recreational and several commercial sectors. Two jurisdictions, Maryland and Virginia, manage this fishery within their jurisdiction. The Potomac River Fisheries Commission manages the Potomac River fisheries. Each jurisdiction has made several (and often different) management decisions in the history of this fishery. There are three sources of commercial fisheries data, being from Virginia, Maryland and the Potomac River Fisheries Commission. The latter data are daily from the Potomac River, but are not used or analysed as a potential source of commercial CPUE in the assessment.

Recommendation 16. Use the daily Potomac data in the assessment, which seem to be very rich and useful.

Analytical approach to adjusting for reporting changes

There have been significant changes to the reporting systems, however the most influential are argued to be those that occurred during 1981. A classic Box-Jenkins time series model is used to analyse these changes. Tests of stationarity in some cases showed that the analyses had to be undertaken on the time series that are in the form of differences. The analysis assumes that current data are more accurate and landings from this period are the most reliable. The effects of this analysis on the Virginia and Maryland estimated commercial landings are significant.

Although the method appears reasonable, the very large effect on the results should be validated. Experience has shown that the time of change is more readily agreed upon the quantum. There are several ways of addressing these issues. These should be prioritised based on ease of implementation.

Recommendation 17. Independently substantiate these adjusted changes to the catch data through interviews of dealers/fishers or investigate whether there is a discrepancy within the spatial landings data as the fishery effort moves over time.

Recommendation 18. Undertake model sensitivity tests with unadjusted catch data and an intermediate value of the reporting change factor.

Recommendation 19. Estimate the reporting change factor (or intervention parameter) internal to the model (best option). If the model is unable to estimate this parameter, then an alternative would be to include the reporting change parameter (with variances) as an input value within the model so sensitivity tests of this parameter can be undertaken and the error in this variable could be included internal to the model.

The different gear types in the fishery are not considered, yet these can be substantially different in location and affect the size and type of crab caught.

Recommendation 20. Investigate the potential of dividing the catch data into gear type or at least undertake a gear analysis.

Effort and catch rate

Neither effort or catch rate data are included in assessment. This is due to the fact that data appear to be inconsistent over time and between jurisdictions.

Recommendation 21. Investigation should be carried out to establish whether it is possible to produce a standardised catch rate dataset that can be used in the assessment, as the index only needs to be representative rather than complete. Also, investigate the use of effort data within the assessment.

Recommendation 22. Given the value of this fishery in the region, there should be more emphasis placed on collecting accurate and complete catch and effort data that are well aligned with the needs of the assessment.

The soft and peeler crab catch is based on landings of animals that have survived being kept in a shedding tank, whereas the animals that died during the process is not included as a component of the catch.

Recommendation 23. Either record the catch of the soft and peeler crab at the point of landing or record the length of time the crabs are kept in the tank and estimate the survival rate of these crabs.

Estimating baywide catches in numbers

The catches are usually reported in bushels and need to be converted to numbers. In recent years, the method of converting to numbers has been to calculate the average carapace width of crabs in the population on an annual and sex specific basis based on the Maryland and Virginia trawl surveys, and then apply to a regression of weight to carapace length based on data from Maryland trawl data pooled between 1994 to 2004. It is unclear (where available) why the distribution itself was not used rather than the average.

Recommendation 24. The conversion of catch in bushels to numbers should use the mean weight from the catch for each year both for the past and the future. This means that the mean size in the catch by year is required.

Recreational harvest

Recreational catch is estimated to be 5.3 to 8.5% of the commercial catch based on surveys in 2001 and 2002. Localised tagging studies pointed to recreational catch being much higher than 8% based on very large return rates of tags from the recreational sector. The uncertainty is whether these studies represent the whole bay. These surveys are the only Baywide surveys and are well analysed and are therefore the only reliable data of value to the assessment.

Recommendation 25. Undertake a baywide recreational survey as the resource is increasing and recreational effort may be increasing. This is a priority.

Recommendation 26. Undertake sensitivity tests for larger recreational catches in the assessment.

ToR: Develop and implement assessment models for the Chesapeake blue crab fisheries. In particular, models that permit estimates of the trends and status of the crab population and fisheries on a sex-specific basis should be evaluated.

The present assessment being reviewed is a two-stage sex disaggregated model that uses three indices of abundance from the surveys, one of which is treated as absolute. The two stages are year 0 and year 1+ animals. About 144 parameters (for the Base Case) are estimated in ADMB. A fundamental positive development of this model is the integration of the reference point calculations internal to the model using per recruit analyses. This means that assumptions in the assessment can be consistently carried into the reference point calculations. This is a major improvement and highly supported. In order for this integration to take place the model has to incorporate a renewal function, in this case a modified form of the Ricker stock recruitment relationship.

This model also differs from its predecessors in that it tracks males and females separately. In light of major recent changes to the fishery management that have changed the sex ratio of the catch, this is an appropriate development on previous models.

There is some controversy concerning the chosen formulation of the stock-recruitment relationship since it implies that the population will remain highly productive while there are any females even at zero male population size i.e. there is not sperm limitation in the population. This form of the relationship is not plausible. There is no information on the lower end of the stock-recruitment curve to assume such a relationship. A sensitivity test of using the traditional approach was requested and provided for comparison and, although the results shown did not completely follow through to the stock status relative to the reference points, it shows very similar results to the Base Case (note these were preliminary runs). This is not unexpected as the model is implemented using a reasonable sex ratio based upon observed data. However, the reality is that this model optimises yield-per-recruit at a sex ratio of zero males. This aspect should be clearly stated in the documentation to highlight that this assessment model should not be used to optimise the sex ratio using the assessment's per recruit analysis. This is an important aspect of the model that should be further justified given the precautionary approach.

Recommendation 27. Undertake full sensitivity tests of the stock-recruitment relationship (following this into the reference point calculations) and also develop a function that captures the principle of what is presently formulated but does not optimise at zero males i.e. incorporate a compensatory function in the stock-recruitment relationship.

In the documentation there was some confusion with regard to how the different indices of abundance were linked to crab biology.

Recommendation 28. Develop a conceptual model which would provide a much better description of the biological timeline and how this connects to the assessment.

In many of the equations in the assessment document (e.g. equation 7,11) both M and F are decremented by an effort related parameter. It would be more appropriate to index M by time of year and F by effort.

Recommendation 29. The decrementation in, for example, equations 7 and 11 etc. should index M with time of year and F by effort.

The partial recruitment parameter in equation 11 captures growth but should also capture selectivity. This confuses what are two separate processes.

Recommendation 30. Split the partial recruitment parameter (that ties selectivity and growth) to include gear specific selectivity. The model should also overtly include selectivity of the gear (including how it changed over time).

The initial conditions of the model use two parameters to set up the numbers matrix. This is not mentioned in the assessment documentation. It is more correct to set up the model such that the initial conditions match the assessment model and would therefore require four (not two) parameters. This is not likely to have a large effect on the model, though.

Recommendation 31. For consistency, set the initial conditions correctly allied with the assessment implementation i.e. from 2 to 4 parameters.

The effective sample size iterative method was used for the two trawl surveys, but given the extensive nature of the dredge survey index, only the calculated variance is used. This seems correctly implemented in the trawl survey cases. If possible, a similar process should be undertaken for the dredge survey.

Recommendation 32. Include the calculation of effective sample size for the winter dredge survey in the likelihood of the assessment.

There should be a bias correction in the stock-recruitment relationship so that the estimated parameters and the per-recruit analysis are consistent with international practice and would therefore not be incorrectly used in other studies.

Recommendation 33. Include the bias correction in the stock-recruitment relationship.

The model consistently fits the trawl survey data poorly. This is in part due to the contradictory nature of the data, but also due to the fact that this assessment is set up as both a process and observation error model. The model tends to treat the poor data fits as observation error and therefore does not capture the high and low recruitment years. There would be some benefit in running a test where the model is formulated as either process or observation error driven.

> **Recommendation 34.** Run sensitivity tests where the model is formulated as either process or observation error driven.

ToR: Examine density-dependent exploitation patterns derived from survey-based and model-based approaches.

Lipcius (Assessment working paper 2) undertook an in-depth analysis of density-dependent exploitation of blue crabs in Chesapeake Bay. The paper presented empirical evidence that exploitation rate varies inversely with population size and that exploitation rate is depensatory. It is unclear how or whether this work is carried into the assessment other than through attempting to choose more precautionary reference points (see below). The use of the values in Lipcius could be considered as a sensitivity test.

ToR: Characterize scientific uncertainty with respect to assessment inputs and stock status.

There were several sensitivity tests undertaken with respect to a) M, b) the degree of partial recruitment, c) estimating male M while fixing the female M to the Base Case value, d) estimating the sex ratio and e) estimating the partial recruitment. It is laudable that the sensitivity table includes the sensitivity of the assessment parameters, but also the stock status relative to the reference points. However, in reality the sensitivity tests should be much more extensive than those provided. Several recommendations are included throughout the text in this document and will not be repeated here.

There are several survey indices within the model and the effect of each individual index is not tested within the model. This should be adopted as standard practice.

> **Recommendation 35.** Undertake a sensitivity test of the influence of the different indices as well as implement the dredge survey as a relative index.

There should have been more sensitivity tests and the documentation should provide confidence intervals on outputs.

> **Recommendation 36.** Run sensitivity tests of different input data and parameters and its effects on reference points. Consider systematic approaches such as FAST or a designed experiment.

In the text, there were some model fit statistics provided, but this was inadequate to fully characterise uncertainty. For example, many of the plots of the indices and their fits did not include residuals. Also, a table of the actual output parameter estimates and variances was not provided.

> **Recommendation 37.** Provide more detailed fit statistics like q-q plots and residual plots for all indices.

Two additional assessment models were provided for review or information – the previous assessment model, but updated with recent data and a production model. As a test of model uncertainty, this is highly unusual and laudable. Given the differences in model structure between the main assessment model and these models, the uncertainty in stock status relative to the reference points cannot be tested. Of the three models tested, the SSCMSA, which is the model used to derive the 2010 reference points, is the most appropriate.

ToR: Evaluate and recommend biological reference points for the Chesapeake Bay blue crab population. The potential for implementing sex-specific reference points should be evaluated.

ToR: Evaluate stock status with respect to reference points.

These two terms of references have been addressed together as they tend to be discussed as a unit.

The derivation of the reference point calculations is correctly implemented in the model even though there is an unusual stock-recruitment relationship. The harvest control rule is based on a mixture of model output to set the values for the reference points, whereas the indices over time are calculated from empirical data. The reason for this formulation is that the assessment is not undertaken each year and as such there was a need to develop a harvest control rule that can be used in the intermediate years.

The exploitation limit reference point (overfishing) is based on the fact that the optimal yield against female age 0+ exploitation fraction is insensitive to the input sex ratio value. The exploitation fraction target reference point is 75% of the limit reference, which is partly based on the method applied by the regional Mid-Atlantic and New England management councils.

The biological target index is based on the absolute abundance of female 1+ crabs from the winter dredge survey. However, in this case the female age 1+ (Spawning) abundance relative to catch in the per-recruit analysis is sensitive to the sex ratio assumed. The assessment team chose a spawning size at the lower end of the range from the sex ratio results of 0.6 to 0.8. The ultimate limit reference point is therefore 50% of the spawning size at maximum sustainable yield, similar to that used in the Federal system.

The basis of choosing reference points traditionally (at least in other countries) is not only the domain of the assessment scientists. In this context, the report should provide more choices of reference points especially the target reference points. Even so, the use of both the Federal reference points system to define the limit reference points and the regional Management Councils' system to define a target reference point is a step forward.

It is important that when different sensitivity tests are undertaken that they are consistent throughout including the empirical indicators. As an example, the test that assumed the winter dredge survey was a relative index of abundance required that the empirical values of the spawning index had to be adjusted appropriately to incorporate the estimated catchability value.

> **Recommendation 38.** Provide more choices of reference points rather than providing single values for each.

The assessment team has chosen a limit reference point value that is not from the precautionary end of the potential range as argued.

Recommendation 39. Provide the option for more precautionary limit reference points.

Changes to the documentation

The documentation describing the model is well explained and set out mathematically, however the following modifications should be incorporated prior to its final delivery:

a. A conceptual model which would provide a much better description of the biological timeline and how this connects to the assessment. This component should also specify the assumptions.
b. More detailed description of previous tests undertaken, provision of estimated parameter values and variances, and uncertainty around model outputs.
c. Sensitivity tests undertaken during the review.
d. The bias correction should be included in the text.
e. Correct the lagged plots as updated during the review.
f. Correct Figure 3.9 to include all the data points.
g. Correct the text when describing the factors chosen in the deltaGLM analyses as presently these are not the best model as highlighted in the tables (Tables 3.4 and 3.5) – the latter was used based on responses to this question during the review.
h. Clarification of the initial conditions as this is not presently documented.
i. Full documentation of the non-precautionary nature of the stock-recruitment relationship particularly at its extremes.
j. The references to the Figures from Figure 6.8 onwards are not always correct which makes confusing reading.
k. The reference to the simple production model Section 6.3.1 is incorrect.
l. Data points on the reference point figures (6.1 and 6.14) should be annotated with year labels, as this would link better with the text.
m. Correct the legend of Figure 6.12.
n. Create a reference point figure which uses the empirical data as analysed by the present 2010 method while using the 2005 reference points i.e. the document would contain three reference points figures as follows: a) old method for standardising the survey data and old reference points, b) new standardised data with old reference points and c) new standardised data with new reference points. The documentation should include an interpretation of the changes when moving from the 2005 to the 2010 method.

Recommendation 40. The assessment document should be updated to include various corrections and clarifications as described in the review report.

Conclusions

This assessment is a valid approach and an improvement on the previous assessments and therefore should be adopted as the basis for management advice. However, the justification for treating the winter dredge survey as an absolute index of abundance needs to be substantiated given that the best model fit is achieved by treating the survey as a relative index of abundance (which is the more widely accepted method) and this has important implications to management. This is the highest short-term priority.

Future work beyond recommendations

This assessment would benefit from being size-based with a shorter time step – this would internalise the growth rate assumptions, allow in-season changes to be accommodated and also incorporate different gear types. Adding some spatial structure to the model should also be considered.

There are clear spatial-temporal changes in the data that have not been fully captured in the assessment. Other crab studies have shown the value of detailed spatial and temporal mining of the data. There would be value in undertaking a detailed and cohesive analysis of spatial and temporal dynamics of different population components on different surveys and their relationship with abiotic factors.

Recommendations with page references

Appendix 1 Documents provided for the review panel

Prior to the meeting of the review panel in Baltimore, the full assessment document was provided to the panel –

Miller, T.J., M.J. Wilberg, A.R. Colton, G.R. Davies, A. Sharov, R.N. Lipcius, G.M. Ralph, E.G. Johnson, and A.G. Kaufman. 2011. Stock assessment of the blue crab in Chesapeake Bay 2011. UMCES Tech. Ser. No. TS-614-11-CBL., University of Maryland Center for Environmental Science Chesapeake Biological Laboratory, Solomons, MD.

Files containing the catch data, control rule calculations, ADMB code and R scripts were provided to the review panel.

In addition four working papers were provided:

- Lipcius, R.N. and A. Smith. 2011. Survival, longevity and natural mortality of mature female blue crabs. Assessment Working Paper 1 to Stock assessment of the blue crab in Chesapeake Bay 2011.

- Lipcius, R.N. 2011. Density dependent exploitation of blue crab in Chesapeake Bay. Assessment Working Paper 2 to Stock assessment of the blue crab in Chesapeake Bay 2011.

- Miller, T.J. 2011. Development and application of simple production models to the Chesapeake Bay blue crab fishery. Assessment Working Paper 3 to Stock assessment of the blue crab in Chesapeake Bay 2011.

- Miller, T.J. 2011. Application and update of a catch, multiple survey model to the Chesapeake Bay blue crab fishery. Assessment Working Paper 4 to Stock assessment of the blue crab in Chesapeake Bay 2011.

Access was also given to past assessments.

Appendix 2 Review Panel Agenda

2010

Blue Crab Stock Assessment Review

Sheraton Baltimore City Center Hotel

101 West Fayette St.

Baltimore, MD

March 29-31, 2011

<u>March 29, 2011</u>

12:30	Welcome & Introductions	Orner
	- Stock Assessment Committee	
	- Review Panel	
12:45	Presentation of the 2010 Blue Crab Stock Assessment	Miller
4:00	General / Open Question Period	Orner
	- Public Comment	
	- Review Panel	
5:30	Adjourn	

<u>March 30, 2011</u>

8:30 Term of Reference Review and Discussion

I. Critically assess and where necessary revise the life history and vital rates of blue crab in the Chesapeake Bay that are relevant to an assessment of the stock.

II. Evaluate and recommend biological reference points for the Chesapeake Bay blue crab population. The potential for implementing sex-specific reference points should be evaluated.

III. Describe and quantify patterns in fishery-independent surveys. Analyses should include evaluation of the impacts of environmental and abiotic factors on survey catches, to maximize the information content of resultant survey time-series.

IV. Describe and quantify patterns in catch and effort by sector and region, including analyses that examine the impacts of reporting changes and trends in CPUE.

12:30 Lunch

1:30 Term of Reference Review and Discussion (continued)

V. Develop and implement assessment models for the Chesapeake blue crab fisheries. In particular, models that permit estimates of the trends and status of the crab population and fisheries on a sex specific basis should be evaluated.

VI. Examine density-dependent exploitation patterns derives from survey-based and model-based approaches.

VII. Characterize scientific uncertainty with respect to assessment inputs and stock status.

VIII. Evaluate stock status with respect to reference points.

5:15 Adjourn

<u>March 31, 2011</u>

9:00 Review Session *[closed-door]*

- Review Panel to discuss assessment methodologies and develop individual opinions.

- Initiate development of summary documents

12:00 Lunch

1:15 Review Session (continued)

4:30 Adjourn

Appendix 3 Statement of Work

External Independent Peer Review by the Center for Independent Experts

Blue Crab Benchmark Stock Assessment - 2010

Scope of Work and CIE Process: The National Marine Fisheries Service's (NMFS) Office of Science and Technology coordinates and manages a contract providing external expertise through the Center for Independent Experts (CIE) to conduct independent peer reviews of NMFS scientific projects. The Statement of Work (SoW) described herein was established by the NMFS Project Contact and Contracting Officer's Technical Representative (COTR), and reviewed by CIE for compliance with their policy for providing independent expertise that can provide impartial and independent peer review without conflicts of interest. CIE reviewers are selected by the CIE Steering Committee and CIE Coordination Team to conduct the independent peer review of NMFS science in compliance the predetermined Terms of Reference (ToRs) of the peer review. Each CIE reviewer is contracted to deliver an independent peer review report to be approved by the CIE Steering Committee and the report is to be formatted with content requirements as specified in **Annex 1**. This SoW describes the work tasks and deliverables of the CIE reviewer for conducting an independent peer review of the following NMFS project. Further information on the CIE process can be obtained from www.ciereviews.org.

Project Description: The blue crab stock has been subject to Baywide stock assessments on two previous occasions. In the years between benchmark assessments, updates on the stock status are provided by the NOAA Chesapeake Bay Office's Chesapeake Bay Stock Assessment Committee. The most recent update concluded that the stock was not overfished and was not then experiencing overfishing. Since the 2005 assessment, the three management jurisdictions have implemented a range of regulatory changes aimed at attaining the target exploitation rate of 46% of the available stock. Thus, it is appropriate that another, Baywide benchmark assessment be conducted. The blue crab resource, specifically for soft and peeler crabs, in Chesapeake Bay has recently been declared a fisheries resource disaster by the Secretary Commerce. In 2009 and 2010, annual updates (not peer-reviewed) have shown slight improvements in the resource. Blue crab is the most important commercial fishery in Chesapeake Bay with annual Baywide landings recently as low as 50 million pounds – roughly 25 million pounds below the long-term average. 2010 predicted landings (if fished at the target exploitation level) could top 100 million pounds. This is obviously a large fluctuation in landings and thus value of the resource to the Bay community.

The first Baywide stock assessment was conducted using a length-based approach to estimate exploitation, and an unweighted average of the four principal fishery-independent surveys to determine abundance. Consequently biological reference points were crude.

In 2001, the technical subcommittee of the Bi-State Blue Crab Advisory Committee (BBCAC) developed a new management framework that relied on exploitation and biomass threshold and target reference points. Threshold reference points were proposed based on a maintaining 10% of the virgin spawning potential and on the lowest observed abundance in the surveys. A target exploitation rate that would lead to an effective doubling of the spawning stock present in 2001 was also selected. The most recent Baywide benchmark assessment for blue crab in the Chesapeake Bay was conducted in 2005. This assessment critically evaluated and revised estimates of the natural mortality rate, the impact of reporting changes on landings estimates, and spawning potential ratio reference points. The 2005 assessment, using data through 2003, recommended adopting the exploitation fraction, defined as the proportion of crabs available at the beginning of the season that are subsequently harvested, in place of less intuitive measures (F) used in previous assessments. Estimates of exploitation fractions were calculated based on the Baywide winter dredge survey (WDS) and within a modified catch-survey analysis that permitted the use of multiple surveys. The approach used in the 2005 assessment was reviewed by a panel of international scientists with expertise in crustacean fisheries who found that it was a substantial improvement over previous assessments. However, the panel also identified issues to be addressed in future assessments. In particular, the panel recommended exploration of the impact of density-dependent processes in life history traits, improvements to the fishery-independent surveys, particularly with regard to catchability, the possibility of developing a sex-specific assessment model and reference points, and a fuller analysis of the impacts of uncertainty on all aspects of the assessment.

The 2010 assessment and targeted research program is a highly collaborative and integrated program to address specific concerns raised by the international review panel from 2005.

The assessment activities are divided into eight specific Terms of Reference (TOR) that were developed based on the review comments received from panel of experts convened to review the 2005 assessment, and from extensive discussion with managers from MDNR, the Potomac River Fisheries Commission and the Virginia Marine Resources Commission, the three relevant management jurisdictions.

NOAA Fisheries is playing a significant role in coordinating disaster assistance to Maryland and Virginia to ensure a sustainable blue crab fishery in Chesapeake Bay. This 2010 Benchmark assessment and research program represents a large investment by NOAA and the state management agencies and should be reviewed internationally.

The Terms of Reference (ToRs) of the peer review are attached in **Annex 2**. The tentative agenda of the panel review meeting is attached in **Annex 3**.

Requirements for CIE Reviewers: Three CIE reviewers shall conduct an impartial and independent peer review in accordance with the SoW and ToRs herein. CIE reviewers shall have working knowledge and recent experience in the application of stock assessment and crustacean fisheries. Each CIE reviewer's duties shall not exceed a maximum of 14 days to complete all work tasks of the peer review described herein.

Location of Peer Review: Each CIE reviewer shall conduct an independent peer review during the panel review meeting scheduled in Baltimore, Maryland during the tentative dates of 29-31 March 2011.

Statement of Tasks: Each CIE reviewers shall complete the following tasks in accordance with the SoW and Schedule of Milestones and Deliverables herein.

<u>Prior to the Peer Review</u>: Upon completion of the CIE reviewer selection by the CIE Steering Committee, the CIE shall provide the CIE reviewer information (full name, title, affiliation, country, address, email) to the COTR, who forwards this information to the NMFS Project Contact no later the date specified in the Schedule of Milestones and Deliverables. The CIE is

responsible for providing the SoW and ToRs to the CIE reviewers. The NMFS Project Contact is responsible for providing the CIE reviewers with the background documents, reports, foreign national security clearance, and other information concerning pertinent meeting arrangements. The NMFS Project Contact is also responsible for providing the Chair a copy of the SoW in advance of the panel review meeting. Any changes to the SoW or ToRs must be made through the COTR prior to the commencement of the peer review.

Foreign National Security Clearance: When CIE reviewers participate during a panel review meeting at a government facility, the NMFS Project Contact is responsible for obtaining the Foreign National Security Clearance approval for CIE reviewers who are non-US citizens. For this reason, the CIE reviewers shall provide requested information (e.g., first and last name, contact information, gender, birth date, passport number, country of passport, travel dates, country of citizenship, country of current residence, and home country) to the NMFS Project Contact for the purpose of their security clearance, and this information shall be submitted at least 30 days before the peer review in accordance with the NOAA Deemed Export Technology Control Program NAO 207-12 regulations available at the Deemed Exports NAO website: http://deemedexports.noaa.gov/sponsor.html).

Pre-review Background Documents: Two weeks before the peer review, the NMFS Project Contact will send (by electronic mail or make available at an FTP site) to the CIE reviewers the necessary background information and reports for the peer review. In the case where the documents need to be mailed, the NMFS Project Contact will consult with the CIE Lead Coordinator on where to send documents. CIE reviewers are responsible only for the pre-review documents that are delivered to the reviewer in accordance to the SoW scheduled deadlines specified herein. The CIE reviewers shall read all documents in preparation for the peer review.

Panel Review Meeting: Each CIE reviewer shall conduct the independent peer review in accordance with the SoW and ToRs, and shall not serve in any other role unless specified herein. **Modifications to the SoW and ToRs can not be made during the peer review, and any SoW or ToRs modifications prior to the peer review shall be approved by the COTR and CIE Lead Coordinator.** Each CIE reviewer shall actively participate in a professional and respectful manner as a member of the meeting review panel, and their peer review tasks shall be focused on the ToRs as specified herein. The NMFS Project Contact is responsible for any facility arrangements (e.g., conference room for panel review meetings or teleconference arrangements). The NMFS Project Contact is responsible for ensuring that the Chair understands the contractual role of the CIE reviewers as specified herein. The CIE Lead Coordinator can contact the Project Contact to confirm any peer review arrangements, including the meeting facility arrangements.

Contract Deliverables - Independent CIE Peer Review Reports: Each CIE reviewer shall complete an independent peer review report in accordance with the SoW. Each CIE reviewer shall complete the independent peer review according to required format and content as described in Annex 1. Each CIE reviewer shall complete the independent peer review addressing each ToR as described in Annex 2.

Other Tasks – Contribution to Summary Report: Each CIE reviewer may assist the Chair of the panel review meeting with contributions to the Summary Report, based on the terms of reference of the review. Each CIE reviewer is not required to reach a consensus, and should provide a brief summary of the reviewer's views on the summary of findings and conclusions reached by the review panel in accordance with the ToRs.

Specific Tasks for CIE Reviewers: The following chronological list of tasks shall be completed by each CIE reviewer in a timely manner as specified in the **Schedule of Milestones and Deliverables.**

1) Conduct necessary pre-review preparations, including the review of background material and reports provided by the NMFS Project Contact in advance of the peer review.
2) Participate during the panel review meeting in the Baltimore, Maryland during the tentative dates of 29-31 March 2011, and conduct an independent peer review in accordance with the ToRs (**Annex 2**).
3) No later than 14 April 2011, submit an independent peer review report addressed to the "Center for Independent Experts," and sent to Mr. Manoj Shivlani, CIE Lead Coordinator, via email to shivlanim@bellsouth.net, and Dr. David Sampson, CIE Regional Coordinator, via email to david.sampson@oregonstate.edu. Each CIE report shall be written using the format and content requirements specified in Annex 1, and address each ToR in **Annex 2**.

Schedule of Milestones and Deliverables: CIE shall complete the tasks and deliverables described in this SoW in accordance with the following schedule.

22 February 2011	CIE sends reviewer contact information to the COTR, who then sends this to the NMFS Project Contact
15 March 2011	NMFS Project Contact sends the CIE Reviewers the pre-review documents
29-31 March 2011	Each reviewer participates and conducts an independent peer review during the panel review meeting
14 April 2011	CIE reviewers submit draft CIE independent peer review reports to the CIE Lead Coordinator and CIE Regional Coordinator
28 April 2011	CIE submits CIE independent peer review reports to the COTR
5 May 2011	The COTR distributes the final CIE reports to the NMFS Project Contact and regional Center Director

Modifications to the Statement of Work: Requests to modify this SoW must be approved by the Contracting Officer at least 15 working days prior to making any permanent substitutions. The Contracting Officer will notify the COTR within 10 working days after receipt of all required information of the decision on substitutions. The COTR can approve changes to the milestone dates, list of pre-review documents, and ToRs within the SoW as long as the role and ability of the CIE reviewers to complete the deliverable in accordance with the SoW is not adversely impacted. The SoW and ToRs shall not be changed once the peer review has begun.

Acceptance of Deliverables: Upon review and acceptance of the CIE independent peer review reports by the CIE Lead Coordinator, Regional Coordinator, and Steering Committee, these reports shall be sent to the COTR for final approval as contract deliverables based on compliance with the SoW and ToRs. As specified in the Schedule of Milestones and

Deliverables, the CIE shall send via e-mail the contract deliverables (CIE independent peer review reports) to the COTR (William Michaels, via William.Michaels@noaa.gov).

Applicable Performance Standards: The contract is successfully completed when the COTR provides final approval of the contract deliverables. The acceptance of the contract deliverables shall be based on three performance standards:

(1) each CIE report shall completed with the format and content in accordance with **Annex 1**,

(2) each CIE report shall address each ToR as specified in **Annex 2**,

(3) the CIE reports shall be delivered in a timely manner as specified in the schedule of milestones and deliverables.

Distribution of Approved Deliverables: Upon acceptance by the COTR, the CIE Lead Coordinator shall send via e-mail the final CIE reports in *.PDF format to the COTR. The COTR will distribute the CIE reports to the NMFS Project Contact and Center Director.

Support Personnel:

William Michaels, Program Manager, COTR
NMFS Office of Science and Technology
1315 East West Hwy, SSMC3, F/ST4, Silver Spring, MD 20910
William.Michaels@noaa.gov Phone: 301-713-2363 ext 136

Manoj Shivlani, CIE Lead Coordinator
Northern Taiga Ventures, Inc.
10600 SW 131st Court, Miami, FL 33186
shivlanim@bellsouth.net Phone: 305-383-4229

Roger W. Peretti, Executive Vice President
Northern Taiga Ventures, Inc. (NTVI)
22375 Broderick Drive, Suite 215, Sterling, VA 20166
RPerretti@ntvifederal.com Phone: 571-223-7717

Key Personnel:

NMFS Project Contact:

Derek M. Orner
NOAA Chesapeake Bay Office,
410 Severn Avenue, Suite 107A,
Annapolis, MD 21403
Derek.orner@noaa.gov
Office: (410) 267-5676 Cell: (410) 570-2268

Annex 1: Format and Contents of CIE Independent Peer Review Report

1. The CIE independent report shall be prefaced with an Executive Summary providing a concise summary of the findings and recommendations, and specify whether the science reviewed is the best scientific information available.

2. The main body of the reviewer report shall consist of a Background, Description of the Individual Reviewer's Role in the Review Activities, Summary of Findings for each ToR in which the weaknesses and strengths are described, and Conclusions and Recommendations in accordance with the ToRs.

   a. Reviewers should describe in their own words the review activities completed during the panel review meeting, including providing a brief summary of findings, of the science, conclusions, and recommendations.

   b. Reviewers should discuss their independent views on each ToR even if these were consistent with those of other panelists, and especially where there were divergent views.

   c. Reviewers should elaborate on any points raised in the Summary Report that they feel might require further clarification.

   d. Reviewers shall provide a critique of the NMFS review process, including suggestions for improvements of both process and products.

   e. The CIE independent report shall be a stand-alone document for others to understand the weaknesses and strengths of the science reviewed, regardless of whether or not they read the summary report. The CIE independent report shall be an independent peer review of each ToRs, and shall not simply repeat the contents of the summary report.

3. The reviewer report shall include the following appendices:

   Appendix 1: Bibliography of materials provided for review
   Appendix 2: A copy of the CIE Statement of Work
   Appendix 3: Panel Membership or other pertinent information from the panel review meeting.

Annex 2: Terms of Reference for the Peer Review

Blue Crab Benchmark Stock Assessment - 2010

The stock assessment review has the following eight specific terms of reference:

a) Critically assess and where necessary revise the life history and vital rates of blue crab in the Chesapeake Bay that are relevant to an assessment of the stock.

b) Evaluate and recommend biological reference points for the Chesapeake Bay blue crab population. The potential for implementing sex-specific reference points should be evaluated.

c) Describe and quantify patterns in fishery-independent surveys. Analyses should include an evaluation of the impacts of environmental and abiotic factors on survey catches, to maximize the information content of resultant survey time series.

d) Describe and quantify patterns in catch, effort and survey-based estimates of exploitation by sector and region, including analyses that examine the impacts of reporting changes and trends in CPUE.

e) Develop and implement assessment models for the Chesapeake blue crab fisheries. In particular, models that permit estimates of the trends and status of the crab population and fisheries on a sex-specific basis should be evaluated.

f) Examine density-dependent exploitation patterns derived from survey-based and model-based approaches.

g) Characterize scientific uncertainty with respect to assessment inputs and stock status.

h) Evaluate stock status with respect to reference points.